Moderator: Jill Darling
October 8, 2019
1:00 pm CT

Coordinator: Welcome and thank you for standing by. Today's call is being recorded. If you have any objections, you may disconnect at this time. All participants are in a listen-only mode until the question-and-answer session of today's conference. At that time, you may press Star and the number 1 on your phone to ask a question. I would now like to turn the conference over to Ms. Jill Darling. Thank you and you may begin.

Jill Darling: Great. Thank you (Tasha). Good morning and good afternoon everyone. I'm Jill Darling from the Office of Communications here at CMS and I will be your moderator today. I'd like to welcome you, (excuse me) to this listening session on the Medicare Severity Diagnosis related group, complication and comorbidity, major complication and comorbidity comprehensive analysis discussed in the fiscal year 2020 Inpatient Prospective Payment System, the IPPS proposed and final rules.

The listening session will include review of the methodology to measure the impact on resource use and will provide an opportunity for CMS to receive public input on this analysis and to address any clarifying questions in order to assist the public in formulating written comments on the current severity level designation (excuse me) for consideration for fiscal year 2021 rulemaking.

I also have one more brief announcement. This listening session is open to everyone, but if you are a member of the press, you may listen in but please refrain from asking questions during the period for comments, (excuse me), any comments of the call. If you have any inquiries, you – please contact CMS at press@CMS.HHS.gov.

And at this time, it's my pleasure to introduce our CMS Deputy Director for the Division of Acute Care, Michelle Hudson, who will be providing opening remarks.

Michele Hudson: Thank you Jill and thank you everyone for joining us for this listening session. As many of you know, CMS has been working with 3M Health Information Systems on our continuing efforts for making refinements to the current MS-DRG's to better recognize severity of illness.

After the determination of the reason for the admission, the most important patient characteristic used in the definition of the DRG's is the presence or absence of a secondary diagnosis that's categorized as a complication or comorbidity, or a CC, or a major complication or comorbidity, an MCC.

For the past few IPPS rulemaking cycles, we've discussed our ongoing efforts to perform a comprehensive severity level designation review for each diagnosis code that we refer to as our CC/MCC analysis.

The goal was to build on the severity DRG work completed for FY 2008 that was subsequently converted to ICD-10, and to perform again a broad-based analysis of the severity levels within the MS-DRG system specific to ICD-10.

The purpose is to increase DRG homogeneity, recognize the impact of varying severity levels on resource consumption and improve payment equity. This way we can better reflect the changes not only in the accuracy and completeness of the ICD-10 coding of the secondary diagnoses, but also in the characteristics of patients admitted to hospitals and the practice patterns within hospitals as well.

As Jill mentioned, in the FY 2020 rulemaking cycle we described our comprehensive CC/MCC analysis and made proposals that would have resulted in a change in severity levels designation for just under 1,500 ICD-10 -CM diagnosis codes. We received many comments on those proposals with the majority of commenters requesting that the adoption of broad-based changes be delayed in order to provide additional time to evaluate those changes, given the broad scope.

After consideration of the public comments in the final rule, we generally did not finalize any of those proposed changes, because planning the adoptions of those comprehensive changes in the severity level designations--not only-- allowed CMS to further examine potential changes in order to make sure that they would appropriately reflect resource use based on the review of more recent data.

It also allows us to take into consideration an opportunity to provide some background to additional – background to interested parties on this methodology that we use, and to obtain some more feedback from stakeholders in the public prior to the development of the FY 2021 rulemaking, and in particular in advance of the upcoming November 1, 2019, deadline for MS-DRG requests for FY 2021.

So with that- as indicated in the announcement for this listening session, CMS has made available an "impact on resource" file that we'll be reviewing as part of the call today to help illustrate that methodology.

With that I'll turn the call to Elizabeth McCullough, Director of the Clinical and Economic Research Group at 3M Health Information Systems, which is when a contractor is responsible for supporting CMS with the DRG

maintenance contract. And she'll talk about the process she used to conduct our comprehensive CC/MCC analysis. Thanks, Liz.

Elizabeth McCullough: Thank you Michelle, and good afternoon everyone. The process that we will be describing here today was used by the team to evaluate each individual diagnosis code and to determine its presence of a secondary diagnosis condition resulting in an increased resource use for hospitals across all DRGs, and designate each code as either a major CC, a CC or a non-CC severity level.

Diagnosis codes classified as a major CC really reflect the highest level of severity. The next level of severity includes diagnosis codes classified as a CC. A CC is really a "non-major CC," but for simplicity's sake we tend to refer to that as just a plain old CC. The lowest severity level is for the non-CCs. Non-CCs are diagnosis codes that do not significantly impact the severity of use of illness or on resource use.

So the categorization of the diagnosis as either a major CC, a CC, or a non-CC was accomplished using an iterative approach in which each diagnosis was evaluated to determine the extent to which its presence as a secondary diagnosis resulted in an increased hospital resource use.

In order to begin this iterative process, it's necessary to start with an initial categorization for each diagnosis code as a major CC, a CC, or a non-CC. For the results that we are reviewing today in the report that was provided, the starting point for this categorization was the MS-DRG Version 35 severity level assignment for each diagnosis code.

So looking at the workbook provided for this meeting, it was really intended to provide additional information-- as Michelle articulated-- on this evaluation

process, provide a report tool used in the evaluation, provide report variable definitions, examples of computations and flagging criteria that was referenced when the evaluating team looked at each secondary diagnosis code.

So let's look at the spreadsheets and the workbook file. The very first worksheet labeled File Description provides an overview of the workbook and content included in the following four worksheets that provides the tabs, the tab descriptions, and on the report tab we'll go through each of the columns, the title columns and a description. We'll get into that a little bit more as we work through this.

The second tab is labeled 1-SDX Codes Impact on Resource Use. This is where you may choose to spend most of your time after this meeting in reviewing the content of the actual report tool used in the evaluation of the secondary diagnosis. But before reviewing the report tool worksheet, one must understand how this report was developed and the meaning and key resource impact variables included in the report.

So let's skip over to the third worksheet labeled, 2-Data Review Examples. I will first walk through aspects of this report development, and then highlight the purpose of the information in the fourth worksheet labeled 3-Flag Criteria, followed by spending most of our time going through some examples that are located on the fifth worksheet labeled 4-Clinical Examples.

So if we all can look at the third worksheet labeled 2-Data Review Example as our next point. So in order to evaluate the effect of each diagnosis code on total resource use for cases in which the diagnosis is a secondary condition, these cases were grouped into three subsets or subgroups, based upon the presence and the severity level of the additional secondary diagnosis codes on the case, as either an MCC, a CC or a non-CC status.

We devised a system that would allow us to measure the value of each secondary diagnosis when it is present in a patient case. So for each secondary diagnosis the impact and resource use was measured by the following three subsets of patients-- and the information if you're on the spreadsheet, I'm looking at Rows 19 through 24.

And we're looking at patients' subset group-- or patients-- or Subset Group 1, really, are patients with no other secondary diagnosis or all of the secondary diagnosis codes are non-CC on the record. Those are cases that are categorized as Subgroup 1.

Subgroup 2 are patients with at least one secondary diagnosis code that is a CC, but there are no secondary diagnosis codes that are a major CC. So all of those patients would be categorized as Subgroup 2. Subgroup 3 are patients with at least one or more major secondary – major CC secondary diagnosis codes.

So for each of these subgroups within a DRG, a national average cost is computed. The costs used in this report were computed using the same charge standardization approach used to compute cost computations in the MS-DRG relative weights. This resulted in a set of expected cost values for each base MS-DRG for each of the three subgroups, from which then an actual cost can be compared using indirect rate standardization techniques and incorporating the DRG for risk stratification in the comparison of the actual cost to the expected.

The actual to expected costs are not compared in your typical A to E ratios that are found on many reports. Instead, a numerical resource impact value

was defined. And if you're on the spreadsheet, I'm going to go over Rows 24 through 30 explaining what the values mean for the resource impact values.

So for an impact value around zero it's meaning that the actual costs were significantly below the expected value for cases that only were in the non-CC subgroup. An impact value around 1.0 means that the actual cost was approximately equal to the expected value for the non-CC subgroup.

An expected – an impact value of 2.0 means that the actual cost is approximately equal to the expected value of the CC subgroup, and an impact value of 3.0 is approximately equal to the expected value for the major CC subgroup.

If you are following along with the worksheet on the number called 2-Data Review Examples, again the values that we just went over and their meanings are shown on Rows 27 through 31.

So for each diagnosis code for which Medicare data was available, the resource team evaluated – looked at these types of reports and evaluated to determine it's – the code's resource impact and use, and determine the most appropriate CC subclass, so either a CC, an MCC or non-CC - a major CC or a non-CC assignment.

But in order to make this determination, the average cost for each subset of cases was compared to the expected cost for each subset. The format shown in the example of Row 11 of the spreadsheet was set up for diagnosis code J44.1. And so I'm going to walk through this example so you get a feel for how to read these reports.

So we show each secondary diagnosis, we give the description, we indicate what the current severity designation level is – it's either a CC, a non-CC or a major CC. For J44.1, the COPD with acute exacerbation, the starting severity level in Version 35 was a CC.

We then show for each of the subgroups, Subgroup 1-- again non-CC, Subgroup 2, the cases with a CC –secondary diagnosis codes, and Subgroup 3, the cases that had at least one or more major CC's, we showed the count of cases again that had that secondary diagnosis code SDX J44.1 that were in either Subgroup 1, Subgroup 2, or Subgroup 3.

And then we also show the impact value. The impact on these reports and in this file is either called C1 or Impact 1, C2 or Impact 2 or C3 or Impact 3, and they are the measure of impact on resource use of patients in each of the subgroups.

So when performing the evaluation of each diagnosis code, we really look for both consistency of the trend of the subclass values across the three subclasses, and the volume of cases were considered in the weight given to the subclass value. Determinations involving the diagnosis or subclass that had low volumes were based upon both not only the cost of the cases, but we also looked at the cost and the volume associated with diagnosis codes that had related conditions.

So if we had a low-volume diagnosis code, we would look at the family of codes that were similar to say COPD clinically, and look at the statistics for those codes that had higher volumes before making a clinical decision about any of the low-volume secondary diagnosis codes.

Several factors were considered in the evaluation-- the diagnosis codes including the clinical aspects, which is most important, and then the cost, and then the coding of the diagnosis. Each diagnosis code was evaluated from a clinical perspective, with the clinical team stating the clinical hypothesis level assignment followed by a statistical review to validate that clinical hypothesis.

In general, current CC designations of a diagnosis was changed if the average cost or values of its subclasses were more-- likely more closely approximated the expected value of another CC subclass, or if the clinical team believes the diagnosis is clinically significant.

So it's important to understand, if the impact value for a diagnosis was closer to 1.0, then it's more likely similar to the expected cost of a non-CC group. If the impact value is closer to 2.0, then it is more likely that the secondary diagnosis code has expected costs more like the CC subgroup. and an impact value closer to 3.0 means that it's more likely acting – the costs are more like the expected of the major CC subgroup.

So let's walk through this example so you can get a feel for it for J44.1. You can see that looking at the stats, there are 272,101 cases in Subgroup 1. That means they have that many cases with secondary diagnosis code J44.1 that had no other major CC or CC secondary diagnosis codes – only non-CCs.

The impact value under Column C1 or Impact 1, shows a value of 1.59. This means that the impact on resource use for the secondary diagnosis is greater than expected – is greater than 1.0 for the non-CCs by an amount that is 59% between the difference between the expected value of the CC and the non-CC. So it's closer to the expected value of the CC, which is a 2.0, and it's higher than what was expected of a non-CC.

So that's how you interpret the impact value, and that makes sense, since the secondary diagnosis code is currently being assigned as a CC, and the impact resource that the secondary diagnosis code J44.1 is having is much higher than that of a non-CC, and it's much closer to that of a CC. So then the clinical teams will look at that, they will make that evaluation and clinically make a decision as to whether or not to keep the level as a CC or upgrade it or downgrade it.

When you look at the Subgroup 2 information, you can see that there's a lot of cases 865,004 cases, that again have that secondary diagnosis code J44.1, but also have one or more additional secondary diagnosis codes that were a CC. The impact value is 2.29.

Again, indicating that with the presence of this secondary diagnosis code J44.1, it is more resource intense than the expected value for CC cases. It's 29% between the difference between the expected value of a major CC and a CC, but is much closer to that of the CC subgroup. So again the teams will look at that and take all of that into consideration as they evaluate.

In looking at the Subgroup 3, there are 369,345 cases again with J44.1 as a secondary diagnosis that have one or more secondary diagnosis codes that are a major CC. The impact value is 3.0. It is exactly the expected value of cases that have a major CC, indicating that J44.1 does not require additional resources above and beyond what is already captured with other cases that have a major CC. So with all that in mind, the clinical team would make a proposal for what the definition for the designation for the CC level should be for that code.

So a lot of information is on this spreadsheet, the actual calculations of the actual to the expected is shown below for your review.

If we move on to the next tab called 3-Flag Criteria, what this is trying to do is provide information detailing out how select diagnosis codes are flagged for additional statistical review beyond the initial clinical review. The volume of cases for each subgroup and the impact value were evaluated statistically based upon the initial non-CC, CC or major CC designation of the secondary diagnosis code.

In looking at the information on the reports, you'll see there is a field called Flag Criteria, and it's designated with a NC of 1+, 2+, 2-, 3-. So I'm going to go over what those values mean so that you can understand how to interpret them when we get to the example report. And again, this is all on the spreadsheet called 3-Flag Criteria and I'm looking at Rows 2 through 6.

So a secondary diagnosis code currently designated as a non-CC that has case volume and a higher than expected impact value were flagged for review as potential candidates for bumping up, and the flag was set to 1+. Secondary diagnosis codes that were designated as a CC that had case volume and higher than expected impact values, were flagged for review as potential candidate for bumping up and the flag was set to 2+.

Secondary diagnosis codes that were designated originally as a CC, with case volumes and lower than expected impact values were flagged for review as a potential candidate for bumping down and the flag was set to 2-. Secondary diagnosis codes designated as a major CC originally that had case volume and lower than expected impact values were flagged for review as a potential candidate for bumping down and the flag was set to 3-.

If none of the criteria was met to set the flag, then the flag was set to the default NC, meaning "no change" could be recommended statistically. This

could be either due to low volume of cases or the impact values are within the range of expected based upon the current non-CC, CC, major CC designation for the secondary diagnosis code.

It doesn't mean that the clinical teams will not review these codes and make a clinical determination, either go with or not go with the statistical, try and stay consistent with the family of codes – that full entire process happens. The statistical flagging is just there for an additional piece of information to take into the broader consideration.

So now let's move on to the last tab in the workbook called 4- Clinical Examples. So here are a set of seven diagnosis codes that we can kind of walk through and give you a feel for ultimately how to go back to the main report tab called 1- SDX Codes Impact On Resource Use, because the format is exactly the same. This is just a small subset that we can walk through and talk through to get familiar with how to again understand and review this report.

So again, we're going to start with Row 2, secondary diagnosis code D68.51, Activated protein C resistance. In Version 35, the starting off level was – this code was set at a CC. You can see that there's plenty of volume in each of the subgroups, so statistically it's strong, and in looking at Subgroup 1 for over 6,000 cases, the impact value is 0.963. It's basically acting like a non-CC when there are no other CC or MCC's on the record.

When you start looking at Subgroup 2, there's a lot more cases-- over 15,000 cases—again, that have the secondary diagnosis code D68.51, and the impact value is 2.027, which means it's acting like a CC. There's already additional CC's on the record, so this particular secondary diagnosis code D68.51 is

really not having a lot more additional impact above and beyond what the other secondary diagnosis codes that are a CC are already having.

And then when you look at Subgroup 3, again a lot of cases, over 11,000, the impact value is 3.071. Again, very close to the expected 3.0. The secondary diagnosis code D68.51 again is really not having any additional resource impact on these major CC cases.

So because of all of that, this case was flagged as a 2-, meaning that it started off as a CC, it's statistically is more like the cases of a non-CC, and was flagged for review for the clinical teams to look at and potentially be a candidate to be downgraded to a non-CC.

If you look at the next row, Row 3, secondary diagnosis code D70.9, Neutropenia unspecified-- again, in each of the subgroups there's plenty of volume. This code currently in Version 35 was a non-CC, and if you look at Subgroup 1, over 3,000 cases, the impact value is 1.737. It is much closer to the expected value of the CC cases – much closer to a 2 than it is to a 1, which is the non-CC cases. So that right there is a potential candidate to say it's acting more like a CC than a non-CC.

And then we look to Subgroup 2, again 20,000 cases, plenty of volume, and the impact 2 shows a value of 2.445. So, above and beyond the impact of the secondary diagnosis, the expected value of 2.0. D70.9 is still having impact on resources above and beyond the other CC's in that subgroup.

And then if we move along to Subgroup 3, again this is where the majority of the cases are, they already have a major CC, 32,000 cases that also have D70.9 and a major CC, the impact is still higher than expected, the actual cost is still higher than expected for these cases in a major CC.

So being the fact this secondary diagnosis code is designated as a non-CC, the statistics indicate that it should be bumped up under clinical review. It does not indicate that it should be a CC or a major CC, it just says it's a candidate for review with a potential for a bump up statistically.

So now if we look at the next code on Row 4, E83.51, Hypocalcemia-- again, in each of the subgroups, plenty of volume. It starts off as a non-CC in Version 35, the impact values are very similar to the code that we just reviewed with neutropenia. So no surprise, the statistics recommend that this particular code be bumped up, and then the clinical team would review that and make a proposal as either a CC or a major CC.

So now let's look at the next two codes on Row 5 and 6. Let's start with Row 5, Z68.41 body mass index. This is-- currently in Version 35 was a CC designation. Again, plenty of volume in each of the subgroups of cases that have Z68.41 as a secondary diagnosis code.

When you look at the impact values, so starting off with Subgroup 1, the Impact 1 value of 1.109 means that the actual is slightly more than the expected value for the non-CC subgroup. So it is having some impact but not very much – not as much, and it's much closer to 1.0 than it is 2.0, the expected value of the CC subgroup.

When you look at Subgroup 2, again these are cases that already have an additional CC, you would expect their subgroup-- their expected value to have around 2.0, and the Impact 2 value for Z68.41 is 2.088. It is pretty much looking at – looking – the actual is very much like the expected value of those cases with the CC, and these cases already have additional CCs. So the body

mass index Z68.41 code is really not having a whole lot more additional impact above and beyond what is already classified as a CC.

And then if we look at Subgroup 3, again over 200,000 cases, the Impact 3 value for Z68.41 is 3.091. Again, it's very close to the expected value for other cases with a major CC of a 3.0. So because of that, and the fact that without any major CC or CC's in Subgroup 1 it is very much acting like a non-CC, the data supports downgrading this code from a CC to a non-CC.

So again, the clinical team would evaluate that, take all of that into consideration, look at all the stats for the body mass index, including the next row, Z68.42, and then look at those statistics and ultimately make a decision on whether or not they want to downgrade these codes from CC to non-CC.

Now let's go to Row 7, secondary diagnosis code B37.81. So again, in each of the subgroups, plenty of volume-- not as much as the other codes, but there's still plenty of volume, over thousands of cases in each subgroup.

In Subgroup 1, where there's-- again, the case has B37.81-- the impact value is 2.195. It is higher than that of the expected cost of the CC subgroup. So it's definitely taking on additional resources and having a bigger impact on resources above and beyond the cases that have a CC. The original designation – the current designation of Version 35, this case is already a CC. So the impact value of 2.195 is reasonable as a starting point as a CC.

However, when you look at the Subgroup 2, almost 12,000 cases have the secondary diagnosis code B37.81 but also has an Impact 2 value of 2.741, meaning that the interaction of the secondary diagnosis B37.81 with other secondary diagnosis codes that are a CC, still has an impact on resource consumption higher than expected, with a value much greater than 2.0, which

is what it would be expected if it had no impact, and it's much like-- more closer to a major CC than a CC above-- beyond what was already recorded with the secondary diagnosis codes of a CC.

And then when you look at Subgroup 3, again over 21,000 cases had that secondary diagnosis code B37.81, as well as a major CC additional secondary diagnosis code. The impact value is 3.551 for Impact 3 Subgroup. The secondary diagnosis code is the actual costs are above the expected costs of that of a major CC. So with all of that in mind, the data supports increasing the secondary diagnosis code designation from a CC to a major CC.

It's not just the Impact 1 score that it is looking like a CC, but it's also the interaction of other secondary diagnosis codes that are a CC or a major CC. So that's why you look at all three subgroups.

The last example I want to walk through is G93.5, secondary diagnosis code Compression of the brain, that's on Row 8. Again you have a good amount of volume of cases. In Impact 1 Subgroup you had almost 4,000 cases, an impact value of 0.924. It's very much like the non-CC cases, and currently this diagnosis code is designated in Version 35 as a major CC.

So the impact value is definitely not in line with that of a major CC, which you would expect the impact value to be around 3.0 as it was-- the resources, the actual resources were that of the expected of a major CC. So that's just a flag right there.

Then we start to look at Subgroup 2, a lot more cases, almost double the amount of cases, almost 8,000 cases have G93.5 plus an additional major – an additional secondary diagnosis code of CC. The impact value is 2.012. It's right on the expected of having a secondary diagnosis code of CC, and

indicating that G93.5 is really not providing any additional resource use consumption above and beyond what is already captured by the other CC's on the record.

If you then look at Subgroup 3, that's where the majority of your cases are, over 34,000 cases. So that's really driving where all the volume of the cases are, and the 34,456 represent the cases again that had G93.5 with an additional secondary diagnosis code that was a major CC, and the impact value was 2.835, which is less than the expected value of cases with a major CC of 3.0.

So that's a lot to take into consideration. So that, as statistically, the secondary diagnosis code is really not the actual-- is definitely less than the expected across-- almost across all of the three subgroups.

The other thing to keep in mind is that on Column K, "% Died," it's also interesting to note why we include this information in the report, for the clinician teams to also see what is the percent died, because that could have an impact on the amount of cost and resources that ultimately would go into the analysis, and you could see that 31.85% of the cases did die. It's a much higher mortality rate.

So that could have an impact on why some of the impact values are slightly lower, and all of that would be taken into consideration as to whether or not the clinical teams want to downgrade the major CC designation to either a CC or a non-CC.

Statistically, the secondary diagnosis code G93.5 is not acting like a major CC, but that also again could be because of the high percentage of patients that die. So the clinical teams may choose to keep it at that higher level-- but

those are the types of discussions that are had and how we can evaluate this type of report.

So I just went through a bunch of examples. I went through a lot, and so you have the workbook so that you can go back through it. Hopefully you took a lot of notes and are ready to jump into the tab called 1-SDX Impact on Resource Use.

So this is a – again it's the same format that we just went over in the clinical examples. You can filter on certain codes, you can filter on the--you know, the CC level for Version 35, you can look at what the flag was indicating. You can look at a variety of-- filter on a variety of items in order to peruse and make your comments on some of the suggested changes that were being proposed.
So hopefully that was informative, and I will send it back to Michelle.

Michele Hudson: Thank you, Liz. That was fantastic. Thanks for walking us through on how to read and interpret these statistical reports and sort of give us a better feel for how we use this information along with our clinical review to come up with our comprehensive analysis.

And I think at this time, we'll just go ahead and we'll send it over to the moderator. We can open up the lines for any general feedback or questions from the participants.

Coordinator: We will now begin our question and answer session. If you would like to ask a question please press Star the number 1 from your phone, unmute your line and speaking clearly when prompted. Your name is required to introduce your question. If you would like to withdraw your question, please press Star and the number 2.

Again if you would like to ask a question please press Star and the number 1. One moment as we reframe the questions. Our first question comes from Dr. Erica Reemer. Your line is open.

(Erica Reemer): Good afternoon. I actually have a few questions, but I'll ask them one at a time and I'll you know let other people go after me. So my first question is, how do you account for common resource utilization? So if a patient has an acute exacerbation of COPD and they have acute and chronic respiratory failure, they may have common resource utilization.

How can you tease that out and see whether the impact is what it would be for some – like I can understand if you have a CC that's completely unrelated then you can see how this particularly affects the resource utilization, but if they have common resource utilization, I don't understand how you could necessarily determine that this – you know it's not necessarily additive, so I don't understand that piece of it.

Elizabeth McCullough: This is Liz McCullough. You're right. It's not exactly additive. You're trying to get at a designation of either CC, non-CC, or MCC. You're not trying to pull out a regression formula with a coefficient value, and that's where that would be a whole separate conversation, and then you have a lot of other challenges with that approach.

So what we're trying to do is create a way to designate the level, and then you re-compute as an iterative process, you would then reestablish all of these variables and all of these statistics to see, with a new system and the diagnosis codes falling into their new designation level, if that diagnosis and the other diagnosis are then meeting the expectations or the expected values.

And at the end of the day, again, it's a clinical decision, so the statistics may not fully tease out ever. Ultimately when there are differences in the evaluation of the statistics and the clinical approach, the teams typically go with the clinical designation, and just simply use the statistics as a way to kind of validate aspects.

Michele Hudson: We'll take our next question please.

Coordinator: Our next question comes from (Gregory Watson). Your line is open.

(Gregory Watson): Thank you. Thank you for having this session, as someone who's tried to replicate this in the past. This is very much appreciated. A couple of questions or requests for the future.

One, is it possible to every year report these thresholds you're using for the different levels? Second is, can you on this tab with-- one where you show the different impacts, is it possible to show the raw dollars that have been computed for each code? That's sometimes useful to see for those of us who are trying to replicate. And the third question is more a request, if you could describe a little bit of, where does the CC exclusion list come in and how is that developed?

Michele Hudson: Hi (Greg). This is Michelle Hudson at CMS. Thank you for those comments. We would be happy to take your request into consideration. We're considering those types of things, as well as making a (workbook) publicly available for in the future.

So if you wouldn't mind, if you could send us those comments to the MS-DRG classification mailbox, and I'll give that email address at the end of the call, that would be greatly appreciated. Liz, I don't know if there is anything

right now that you can speak to about the CC exclusions list, or if that's something better handled through the mailbox?

Elizabeth McCullough: As far as this report is concerned, if you're asking about the CC exclusion list and how that applies to this report, any secondary diagnosis code that is excluded due to the principal diagnosis would not be evaluated in this report, it would be excluded.

(Gregory Watson):   Thank you.

Coordinator: Our next question comes from Karen Heller. Your line is open.

Karen Heller: Hi, and hi (Greg Watson). So, the question we struggled with during the comment period had to do with the interaction-- in other words, when you decide to downgrade a CC or, you know, make it no longer a CC or a MCC, then those cases change subclasses, and how do you consider the interactive effect of dropping certain codes as you evaluate each subsequent code?

Do you know what I mean? In other words, something that looks like it's not different from others in the subclass might look different if the subclass changes.

Elizabeth McCullough: Absolutely. And as part of this process, we don't just go through this once, we'll then simulate the impact of making these changes. Once each of the diagnosis codes were evaluated, the ones that were revised, these computations are re-estimated and the computational values for each secondary diagnosis code that had changes with the list, so you know that where on the CC or major CC changes were made, they were all revised and then reevaluated, relooked at.

You know, we went through a lot of effort to document aspects of statistically met criteria, clinically made a decision, all of that, so when it's redone and reevaluated and re-reviewed, we would look especially at those that changed to make sure that nothing else either popped up because other secondary diagnosis codes moved, and then therefore it would move groups of cases from one subgroup to another.

The expected's would change, and all of that would get reevaluated. So it's an iterative process until it was stable.

Karen Heller: And is--in general how many iterations is required?

Elizabeth McCullough: It really just depends on the codes. It really just depends.

Karen Heller: I see. Well, you know, any additional information you could provide about that, you know, in terms of letting us see different versions of it would be very helpful.

Another – my second question, though, is just doing a quick and dirty, you know, analysis of the new table that you provided, it does seem like you're providing the outcome for all of the codes with at least 200 cases – well actually no there's – for all the codes period, but that's one of the things that we ask for. Is that correct?

Michele Hudson: Yes, that's correct.

Karen Heller: Okay, and then it looks like the number of cases with a suggested change is much lower. Do you agree with that, and also that there are fewer downgrades or more upgrades? Can you just comment, you know, on where your current thinking is vis-á-vis the proposed rule?

Michele Hudson: So, in preparation for this listening session, we took into consideration some of the public feedback that we had gotten requesting this additional information, and so we did, you know, put out the additional information to help explain the methodology, and to give folks a better sense of how we use it in our decision making process.

But part of – the purpose of this listening session was to help us get a little bit more feedback from all of the public and the stakeholders in order to prepare for the 2021 rulemaking. And so at this time, we really haven't started revisiting that process yet. We're still collecting and gathering information.

Karen Heller: So we shouldn't interpret the information in this file as suggesting anything for the 2021 proposed rule?

Michele Hudson: Yes, that's correct. Could we go to our next question please since we're getting a little bit short on time? Thank you.

Coordinator: Our next question comes from (Valerie Winkle). Your line is open.

(Valerie Winkle):   Thank you. I have a question with regard to the flag criteria and the volume of cases, because like the prior comments are – I'm looking in your second tab, the 1-SDX Codes, I see volume is very low, but I do see what appears to be a recommended status change from a category.

And so when you talked about the family of codes, did you ever consider aggregating across the family, like the neoplasms, for example, and seeing if the results changed?

Elizabeth McCullough: We did not aggregate across family of codes. The clinical teams upon their review would look at the whole family of codes, put together their principles from which-- and discuss clinically the aspects of these secondary diagnosis codes and the impact they are having, and then ultimately make decisions to change a designation or not.

The flagging criteria again is purely statistical. No change simply means that it didn't hit the flagging criteria to recommend an increase or a decrease. It could be because there was just not enough volume, low volume, so they all just said "no change."

We could probably improve the flagging criteria to get out of – to put LV for low volume or something else so that "no change" truly recommended that statistically there was volume, but statistically it is good with the current designation so that's probably an improvement that we can make. So I just want to make sure everyone is aware of that.

The other thing that I just want to also clarify is that on the report you'll see in Column C and D, if the diagnosis code is new, it has a N. Obviously there's not going to be any data to support it, but the old deleted code is probably in the data. So we will also then reference for the new code what the deleted mapped code is so that the clinical teams can then look back at that mapped code, look at those statistics and then ultimately make a decision as to how to apply that information, those statistics, forward to the new codes, and decide clinically what they want to do with the designation.

(Valerie Winkle): Okay.

Coordinator: Our next question comes from (William Haik). Your line is open.

(William Haik):   Thank you. I couldn't understand from the Federal Register that the proposed

                    note, how moderate malnutrition could have a higher category than severe

                    malnutrition. Was that just an error or how could that be a result of an

                    iterative clinical process?

Michele Hudson: Hi. Thank you for that question. That was similar to some of the comments

                    that we received on the proposed rule. You know the purpose of this call was

                    really just to sort of give a broad overview of the process and we can-- for any

                    detailed comments, we encourage you to please put those in writing to our

                    MS-DRG change request mailbox, and I'll be giving that mailbox at the end

                    of the call and then we can take that feedback into consideration. Thank you.

(William Haik):   Thank you.

Coordinator: Our next question comes from (Mark Hartstein). Your line is open.

(Mark Hartstein): Hi. Hey Michelle. Hey Liz. It's good to hear your voice. I think I recall

                    going through an exercise similar to this probably about 12 years ago for the

                    MS-DRGs. I see on Tab 1-SDX codes there are 64,884 lines in that

                    spreadsheet. So I presume what you would do is only evaluate and provide

                    clinical oversight to the ones where there is a change or recommendation

                    going up or going down?

Elizabeth McCullough: So the clinical team would take an evaluation not only of those that

                    showed a flag indicated statistically, but also would look down and clinically

                    look at family of codes to do that analysis as well. And so, yes, it's a lot of

                    work to go through individual-- code by code by code-- and for the most part,

                    the clinical teams --the family of codes – the broad set of codes, but absolutely

                    they would evaluate every single code that had a flag indicating statistically to

                    make a change, but those are not the only codes that were looked at.

Again, a broad family of codes, to make sure there was consistency as well as they could across the codes, and the principles around making those decisions.

(Mark Hartstein): I know this was done for ICD-9, because I was involved in that.

Elizabeth McCullough: Yes.

(Mark Hartstein): This is the first time this is being done with ICD-10?

Elizabeth McCullough: That is correct.

(Mark Hartstein): I see. So the current classifications are based on the crosswalks between ICD-9 and ICD-10?

Elizabeth McCullough: That is correct.

(Mark Hartstein): I see. All right. Thank you.

Michele Hudson: (Tasha), we'll take our last question please.

Coordinator: Okay. Our last question comes from (Gregory Watson). Your line is open.

(Gregory Watson):   Thank you for taking my next question, which is prompted by something that (Valerie) said. In your methodology, on the flag criteria tab, it depends on the volume being greater than 200. We have greater granularity in the ICD-10 than ICD-9. Did you adjust the threshold when moving to ICD-10, or did everything stay the same? Just wondering if some things might not be picked up because of the greater granularity.

Elizabeth McCullough: So when this research was done back in 2007-- and yes, I was a participant in those meetings as well-- we really didn't have official flagging criteria. It was literally large, long meetings with rooms of clinicians going through every single code row by row, one by one, looking at the family of codes, looking at all of the information. There was no hard, fast kind of suggested rule criteria at that moment.

So what the analysts did who have been working on these types of projects for the last 40 years-- not myself, only the last 30 years for myself-- we came up with this general flagging criteria. It's kind of like the white flag, it's not the red flag, stop, do nothing, don't you know – stop, don't pass go kind of thing.

It's the white flag that says, hey, this is something to look at, and the teams kind of came up with around 200 cases was the right amount for this type of analysis – again, it's just a flag. The teams look at things, they step back, they look at the family of codes, and then they look at everything around that to ultimately make their decision.

(Gregory Watson):   Thank you.

Elizabeth McCullough: Mm-hm.

Michele Hudson: So, thank you everyone for joining us for this listening session. This dialogue has been really valuable to the CMS team, and we look forward to continuing this dialogue with you all.

And I apologize if we weren't able to get to your question, but I'd please like to remind you as was indicated in the announcement, that if you have any comments or feedback that you would like us to consider in advance of the FY 2021 rulemaking, please send those comments to our MS-DRG classification

change request mailbox. That email address is MSDRGclassificationchange, all one word, at CMS.HHS.gov.

And as a reminder, in order for those requests to be ensured that they're considered, please submit those requests to the mailbox by the November 1, 2019 deadline, and also if you have any questions that we weren't able to get to today, we'll do our best to respond to those through the mailbox as well. So thank you again, everyone, for joining us, and have a great rest of your day.

Coordinator: Thank you for your participation in today's conference, you may disconnect at this time.

END